

EXHIBIT 3

PLAINTIFFS' SUPPLEMENTAL RFP
TO DEFENDANT GOOGLE LLC RE:
SOURCE CODE

PROPOUNDING PARTIES:

Consolidated Plaintiffs

RESPONDING PARTY:

Defendant Google LLC

SET NUMBER:

Third

Pursuant to Rules 26 and 34 of the Federal Rules of Civil Procedure, Plaintiffs in the above-titled action hereby request that Defendant Google LLC produce to Consolidated Plaintiffs all Documents, electronically stored information, and things that are in Your possession, custody, or control responsive to the following Requests for Production.

DEFINITIONS

1. Each of these definitions and instructions is incorporated into each of the requests to which it pertains. Notwithstanding any definition below, each word, term, or phrase used in these requests for production is intended to have the broadest meaning permitted under the Federal Rules of Civil Procedure.

2. The terms “relating to,” “related to,” “relate to,” “refer to,” or “concerning” or any variations thereof, mean referring to, pertaining to, reflecting, containing, evidencing, supporting, describing, establishing, showing, disclosing, explaining, mentioning, analyzing, constituting, comprising, setting forth, summarizing, characterizing, and/or contradicting, either directly or indirectly, in full or in part.

3. “All,” “or,” and “and” should be understood to include and encompass “any”; “or” should be understood to include and encompass “and”; and “and” should be understood to include and encompass “or.”

4. “Action” means this consolidated action, captioned *In Re Google Generative AI Litigation*, Master Case No. 5:23-cv-03440-EKL (N.D. Cal.)

5. “Agent” means any employee, officer, manager, director, attorney, representative, independent contractor, subcontractor, day laborer, temporary worker, or any other person—regardless of formality of relationship—acting at Your direction or on Your behalf.

6. “AI-Powered Products” means Google products that integrate and/or are powered by one or more Generative AI Models.

7. “Alphabet” means Defendant Alphabet Inc., together with any of its corporations, businesses, subsidiaries, divisions, subdivisions, predecessors, successors, parents, and affiliates, and their respective officers, directors, employees, partners, representatives, Agents, attorneys, accountants, or other

1 persons occupying similar positions or performing similar functions.

2 8. “And” and “or” shall be construed conjunctively or disjunctively as necessary to make the
3 Request inclusive rather than exclusive.

4 9. “Artificial Intelligence” or “AI” engineered system that generates outputs such as content,
5 forecasts, recommendations, or decisions for a given set of human-defined objectives.

6 10. Unless otherwise specified in a Request, “Bard” refers to the large language model (“LLM”)
7 (i.e., AI software designed to emit convincingly naturalistic text outputs in response to user prompts)
8 created, maintained, distributed, and/or sold by Google, including all versions, iterations, and/or model
9 family members of Bard

10 11. “C4” means the dataset created by Google and referenced in the February 10, 2022 research
11 paper entitled LAMDA: Language Models for Dialog Applications.

12 12. “Common Crawl” means the nonprofit organization that crawls the web and makes its
13 archives and datasets publicly available.

14 13. “Communication” means every exchange of information of any nature, whether oral,
15 written, or recorded, from one person or entity to another, and any evidence of such exchange, Including,
16 but not limited to, any correspondence, memorandum, electronic mail messages, social media postings,
17 messages sent or received through social media, collaboration platforms or other online service such as
18 Slack, Workplace, Teams, message board postings, text messages, letters, notes or logs of meetings,
19 diaries, daily calendars, videos, audio recordings, or other records of exchanges between people or entities.

20 14. “Consolidated Plaintiffs” means Plaintiffs Steve Almond, Sarah Andersen, Burl Barer,
21 Jessica Fink, Kirsten Hubbard, Hope Larson, Mike Lemos, Jill Leovy, Connie McLennan, and Jingna
22 Zhang.

23 15. “Document” means any document or electronically stored information as described in
24 Federal Rule of Civil Procedure 34(a) Including, but not limited to, any handwriting, typewriting, printing,
25 photostating, photographing, photocopying, transmitting by electronic mail or facsimile, and every other
26 means of recording upon any tangible thing, any form of Communication or representation, Including
27 letters, words, pictures, sounds, or symbols, or combinations thereof, and any record thereby created,
28 regardless of the manner in which the record has been stored.

This Includes:

- The originals, drafts and All non-identical copies thereof, whether different from the original by reason of any notation made on such copies or otherwise;
- Booklets, brochures, pamphlets, circulars, notices, periodicals, papers, contracts, agreements, photographs, minutes, memoranda, messages, appraisals, analyses, reports, financial calculations and representations, invoices, accounting and diary entries, inventory sheets, diaries, appointment books or calendars, teletypes, telefaxes, thermafaxes, ledgers, trial balances, correspondence, telegrams, press releases, advertisements, notes, working papers, drawings, schedules, tabulations, projections, information or programs stored in a computer (whether or not ever printed out or displayed), and All drafts, alterations, modifications, changes or amendments of any of the foregoing;
- Graphic or aural representations of any kind, Including, without limitation, photographs, microfiche, microfilm, videotapes, recordings, drawing, charts and motion pictures;
- All letters, words, pictures, sounds, or symbols, or combinations thereof stored in or on any electronic, mechanical, magnetic, or optical device Including, but not limited to: (i) computer data storage devices (servers, laptops hard-drives, flash drives, discs, magnetic cards, and the like), (ii) the internet or “the Cloud” (such as e-mail, web posts, social media posts, internet pages, or Google Suite documents (e.g., Google Docs or Google Sheets) etc.), and (iii) information stored on cell phones.

16. Unless otherwise specified in a Request, “Gemini” refers to the multimodal large language model (see *infra*) created, maintained, distributed, and/or sold by Google, including all versions, iterations, and/or model family members of Gemini such as Gemini 1.0, Gemini 1.0 Pro, Gemini 1.0 Ultra, Gemini 1.0 Nano, Gemini 1.5 Pro, Gemini 1.5 Flash, Gemini 1.5 Flash-8B, Gemini 2.0, Gemini 2.0 Flash, and the Gemini App.

17. “Generative AI Models” means all generative AI models created, maintained, distributed, and/or sold by Google LLC other than Imagen, Bard, and Gemini, including but not limited to, all versions, iterations, and/or model family members of PaLM, GLaM, LaMDA, Codey, Chirp, Veo, MedLM, LearnLM, SecLM, Gemma, Code Gemma, Recurrent Gemma, PaliGemma.

1 18. “Google” means Defendant Google LLC, together with any of its corporations, businesses,
2 subsidiaries, divisions, subdivisions, predecessors, successors, parents, and affiliates, and their respective
3 officers, directors, employees, partners, representatives, Agents, attorneys, accountants, or other persons
4 occupying similar positions or performing similar functions.

5 19. “Identify,” with respect to a Document means, to give, to the extent known, the (i) type of
6 Document (*i.e.*, correspondence, memorandum, business record, etc.); (ii) general subject matter; (iii) date
7 of the Document; (iv) author(s) or preparer(s), addressee(s), and recipient(s) of each such Document; (v)
8 the present location of each such Document or copies thereof; and (vi) if applicable, the Bates number of
9 the Document. “Identify,” with respect to an entity, governmental body, department, or agency (not a
10 natural Person), means to state the full name, address, and telephone number for its principal place of
11 business. “Identify,” with respect to a natural Person means, to give, to the extent known, the (i) Person’s
12 full name and title; (ii) current or last known address; (iii) current or last known business and telephone
13 number(s); and (iv) current or last known place of employment. Once a Person has been identified in
14 accordance with this paragraph, only the name of that Person need be listed in response to subsequent
15 discovery requesting the identification of such Person.

16 20. “Including, “including, “include,” or “includes” is used to provide examples of certain types
17 of information and should not be construed as limiting a request or definition in any way. The term
18 “Including, ” “including,” “include,” or “includes” shall be construed as if followed by the phrase “but not
19 limited to.”

20 21. “Information” or “information” means written, recorded, graphic, or other electronic data,
21 including metadata, of any nature whatsoever, regardless of how recorded, and whether original or copy.

22 22. Unless otherwise specified in a Request, “Imagen” refers to the text-to-image diffusion
23 model (*i.e.*, an AI software product designed to generate an image in response to a text prompt) created,
24 maintained, distributed, and/or sold by Google, including all versions, iterations, and/or model family
25 members of Imagen.

26 23. “Image-Text Pair” means a dataset element comprising both an image and its corresponding
27 textual description, used to train AI models on understanding and generating content that combines visual
28 and textual information.

24. “Infiniset” means the dataset created by Google and referenced in the February 10, 2022 research paper entitled *LAMDA: Language Models for Dialog Applications*.

25. “LAION” means Large-Scale Artificial Intelligence Open Network (an organization based in Hamburg, Germany), its predecessor and successor entities, its officers, directors, engineers, researchers, employees, agents, attorneys, representatives, and other persons acting or authorized to act on its behalf.

26. “LAION Datasets” includes all datasets publicly released or privately developed by LAION, whether alone or in collaboration with others.

27. “LLM” or “large language model” means an AI system trained on large amounts of text data to understand and generate human-like language.

28. “MLLM” or “multimodal large language model” means an LLM that accepts input and generates output in multiple modalities and is trained using text, image, audio, and video data.

29. “Person” means any natural persons, proprietorships, legal or governmental entities, corporations, partnerships, trusts, joint ventures, groups, associations, or organizations. Any reference to a “person” shall mean that “person” and all affiliates, divisions, controlled companies, subsidiaries, or otherwise related entities and all of his, her, or its current and former officers, directors, employees, agents, representatives, attorneys, and accountants.

30. “Plaintiffs” refers to the plaintiffs in the above-entitled action.

31. “Source Code” includes computer code and associated comments and revision histories, formulas, engineering specifications, or schematics that define or otherwise describe in detail the algorithms or structure of software or hardware designs. A document that merely quotes or cites filenames, line numbers, directory names, module names, class names, parameter names, variable names, function names, method names and/or procedure names (e.g., for notetaking or identification purposes) without reproducing any lines of source code is not a Source Code Document.” “Such as” or “such as” is used to provide examples of certain types of information and should not be construed as limiting a request or definition in any way. The term “Such as” or “such as” shall be construed as if followed by the phrase “but not limited to.”

32. “Such as” or “such as” is used to provide examples of certain types of information and should not be construed as limiting a request or definition in any way. The term “Such as” or “such as” shall be construed as if followed by the phrase “but not limited to.”

33. “Targeted Models” means Gemini 2.5 Pro; Gemini 2.5 Flash; Gemini 2.0 Pro; Gemini 2.0 Flash; Gemini 2.0 Flash Thinking; Gemini 2.0 Flash-Lite; Gemma 3-27b; LaMDA; GLaM; PaLM (1 and 2); and Imagen (1, 2, and 3).

34. “Third-Party Subpoena(s)” mean and include each subpoena that has been issued or will be issued by You in this case under Rule 45 of the Federal Rules of Civil Procedure.

35. “Training Data” refers to any data used to develop, pre-train, train, test, fine-tune, or modify any AI, machine learning, or deep learning model, Including any textual, visual, audio, or multimedia data, or metadata related thereto. “Training Data” Includes All datasets, structured and unstructured data, images, audio-visual works, textual descriptions, and metadata whether obtained through scraping, licensed sources, third-party datasets (Including LAION-400M), internal datasets, or any other source. To the extent there are multiple variations, transformations, and/or iterations of such data, “Training Data” refers to each separate variation, transformation, and iteration of such data.

36. “You” or “Your” refers to and Includes the specific Defendant producing Documents in response to these Requests, its employees, agents, attorneys, accountants, representatives, predecessors or successors-in-interest, any corporation or partnership under its direction, and any other person or entity acting on its behalf or under its control.

INSTRUCTIONS

1. If, in responding to these requests, You encounter ambiguities when construing a question, instruction, or definition, Your answer shall set forth the matter deemed ambiguous and the construction used in answering.

2. Unless otherwise specified, Documents (Including electronically stored information) shall be produced consistent with and in the manner set forth in the applicable ESI Protocol or, to the extent such ESI Protocol has not yet been entered by the Court, the portions of the ESI Protocol upon which the parties agree.

3. All Documents retained in a machine-readable format should be produced in that machine-readable format.

4. These Requests for Productions should be deemed continuing such that if Your directors, officers, employees, agents, representatives, or any person acting on Your behalf, subsequently discover or

1 obtain possession, custody, or control of any Document or ESI previously requested or required to be
2 produced, and supplemental productions should be provided as additional Documents become available.

3 5. If You claim You are unable to produce a Document, you must state whether that inability
4 is because the Document never existed; has been destroyed, lost, misplaced or stolen; or has never been or
5 is no longer in your possession, custody or control. Such a statement must further set forth the name and
6 address of any person or entity that you know or believe to have possession, custody or control of that item
7 or category of item. If any Document responsive to a request has been destroyed, produce All Documents
8 describing or referencing (1) the contents of the lost or destroyed Document; (2) All locations in which any
9 copy of the lost or destroyed Document had been maintained; (3) the date of any such loss or destruction
10 to the extent known; (4) the name of each person who ordered, authorized and carried out the destruction
11 of any lost or destroyed Document; (5) All Document retention and destruction policies in effect at the time
12 any requested Document was destroyed; and (6) All efforts made to locate any responsive Document
13 alleged to have been lost or destroyed.

14 6. If You object to any item or category of item, Your response shall (a) identify with
15 particularity each Document or thing to which the objection is made and (b) set forth clearly the extent of,
16 and specific ground for, the objection; and YOU should respond to the Request to the extent it is not
17 objectionable.

18 7. If You object that a Document is covered by the attorney-client or other privilege, or is work-
19 product, You shall provide a Privilege Log containing: (1) the name of the Document; (2) the name and
20 address of the person(s) who prepared it; (3) the person(s) to whom it was directed or circulated; (4) the
21 date on which the Document was prepared or transmitted; (5) the name and address of the person(s) now
22 in possession of the Document; (6) the description of the subject matter of the Document; and (7) the
23 specific nature of the privilege claimed, Including the reasons and each and every fact supporting the
24 withholding, and legal basis sufficient to determine whether the claim of privilege is valid with respect to
25 the Document (without revealing privileged information).

26 8. In responding to these Requests for Production, You are to Include Documents (1) obtained
27 from witnesses who gave information to any governmental agency or investigatory body, Including
28 Congress; (2) that constitute, refer or relate to summaries of testimony or other statements in connection

1 with any governmental agency or investigatory body proceeding or investigation, Including before
2 Congress; or (3) obtained on Your behalf in preparation for testimony or interviews before any
3 governmental agency or investigatory body, Including Congress.

4 9. Unless otherwise specified, the Relevant Time Period is from January 1, 2017 until the
5 Present.

6 **REQUESTS FOR PRODUCTION**

7 **REQUEST FOR PRODUCTION NO. 65:**

8 All Source Code for the Targeted Models related to the following functionalities:

- 9 (a) scrapers or crawlers used to assemble training data, including evidence of domains or URLs
10 scraped;
- 11 (b) processing text, images, and image-text pairs, including scripts for dataset assembly and
12 annotation;
- 13 (c) pretraining, fine-tuning, and evaluation, including any scripts used to check for
14 regurgitation or leakage; and
- 15 (d) data infrastructure, including scripts organizing or storing datasets, and any Source Code
16 used to create synthetic or bespoke datasets.
- 17
18
19
20
21
22
23
24
25
26
27
28